

Estimation in Two-stage Adaptive Designs

Per Broberg* and Frank Miller†

February 1, 2016

Abstract

We consider conditional estimation in two-stage sample size adjustable designs and the following bias. More specifically, we consider a design which permits raising the sample size when interim results look rather promising, and, which keeps the originally planned sample size when results look very promising. The estimation procedures reported comprise the unconditional maximum likelihood, the conditionally unbiased Rao-Blackwell estimator, the conditional median unbiased estimator, and the conditional maximum likelihood with and without bias correction. We compare these estimators based on analytical results and by a simulation study. We show in a real clinical trial setting how they can be applied.

Keywords: Adaptive design; Conditional estimation; Sample size recalculation; Two-stage design

1 Introduction

Adaptive design has made its way into the repertoire of practical clinical biostatistics as shown by the discussion [10] as well as the fact that European and US guidance documents for the use of adaptive designs in clinical trials had been developed. Adaptive designs are applied in the confirmatory context (Phase III) as well as in earlier development. For example, recently adaptive Phase II or seamless Phase II/III case studies have been described [19, 8].

In this article, we consider recalculation of the sample size based on an interim effect estimate. For this adaptive design, statistically valid methods are required for inference. Much work has been done in the past decades to provide significance tests which adequately control the type I error rate in the context of sample size recalculation. For example, conditions have been identified [24, 5, 18, 2] where the conventional significance test (which would be done in a non-adaptive situation) is still valid after sample size recalculation in controlling the type I error rate. In these references, the general conditional error approach is applied which says that adaptations are allowed as long as the conditional error rate of the final test does not increase by the adaptation, see [20, 21]. Alternatively, if it is not desired to adhere to these conditions, several good methods exist indicating how modified significance tests can be performed in this adaptive setting with type I error rate control.

However, point and interval estimates are usually of high importance for clinical interpretation of trial results. Therefore it has been identified that inference in adaptive designs needs to focus more on estimation.

The current article will discuss the implications for bias and variance of parameter estimates in scenarios involving two-stage adaptive designs with sample size recalculation. Especially, we will consider conditional inference. [14] exhibit conditionally unbiased estimators for the case of a study with possible futility stop. [13] have for a multi-arm study considered additionally the possibility to remove some of the treatments in an interim analysis. The reference [17] derives conditional maximum likelihood estimators in a group sequential setting.

For the case of a two-stage adaptive design with sample size recalculation, we will develop here conditional estimation methods: the conditionally unbiased Rao-Blackwell estimator, the conditional median unbiased estimator, and the conditional maximum likelihood with and without bias correction.

*Division of Cancer Epidemiology, Department of Clinical Sciences Lund, Lund University, Skane University Hospital, Lund, Sweden. Email: per.broberg@med.lu.se

†Department of Statistics, Stockholm University, 10691 Stockholm, Sweden. Email: frank.miller@stat.su.se

We compare these estimators and the unconditional maximum likelihood estimator based on analytical results and by a simulation study.

For a previously conducted randomized controlled clinical trial which investigated a treatment for schizophrenia patients, we show how the considered estimators can be applied. In this example we show also that our investigations can be applied in contexts which seem slightly different to those formally discussed.

In this article, we will consider the case of normally distributed data with known variance. In Section 2, we provide the general assumptions and the considered sample size recalculation rule, discuss bias of the conventional maximum likelihood estimate (MLE) and mention unconditionally unbiased estimates. In Section 3, we derive several conditional estimators and compare them algebraically. In Section 4, we compare the estimators in a simulation study, show the clinical trial application in Section 5 and conclude with a discussion in Section 6.

2 Background

2.1 General setting and bias

We consider the one-sample case with independent normally distributed observations. Let the observations be $X_1, X_2, \dots \sim N(\mu, \sigma^2)$ with σ^2 known. An interim analysis is performed after n_1 observations (Stage 1) and based on the results, the sample size N_2 for Stage 2 is determined; the total sample size is $N = n_1 + N_2$ (note that N_2 and N are random variables). We use the simplified notation $Y_1 = \sum_{i=1}^{n_1} X_i/n_1$, $\sigma_1 = \sigma/\sqrt{n_1}$; $Y_2 = \sum_{i=n_1+1}^N X_i/N_2$ represents the additional information collected after the interim analysis, $\sigma_2 = \sigma/\sqrt{N_2}$.

A more general set-up, including two-arm parallel group studies with known common variance and observations $\{(X_{1,1}, \dots, X_{1,n}), (X_{2,1}, \dots, X_{2,n})\}$, is touched upon in the Discussion. For now, please note that forming paired differences $X_i = X_{1,i} - X_{2,i}$ will turn that design into the present set-up.

In a study with the aim to test $H_0 : \mu = 0$ versus $H_1 : \mu > 0$, the parameter μ should be estimated by a $\hat{\mu}$ after all N observations. The mean of all observations is the unconditional maximum likelihood estimate (MLE) for μ . Therefore, we write $\hat{\mu}_{ML} = Y = \sum_{i=1}^N X_i/N$ and further, $\sigma_0 = \sigma/\sqrt{N}$.

In [1] it is shown that estimators of the form $\hat{\mu}_{WM,w} = w(Y_1)Y_1 + (1 - w(Y_1))Y_2$, i.e. a weighted mean of Y_1 and Y_2 where the weight w may depend on the first stage outcome, have bias $Cov[w(Y_1), Y_1]$.

To choose the 2nd stage sample size $N_2 = N_2(Y_1)$ as decreasing function of $Y_1 > 0$ is a natural choice as larger sample sizes are needed if the true $\mu > 0$ is small. If one applies equal weight to all observations in $\hat{\mu}_{WM,w}$, i.e. $w(Y_1) = n_1/(n_1 + N_2(Y_1))$, the result of [1] shows therefore that $\hat{\mu}_{WM,w}$ overestimates μ under $\mu > 0$.

In this paper, we consider the following design: We decide in the interim analysis to stop for futility if the observed effect is low (decision $R = 0$ and final sample size $n_f \geq n_1$), to continue to an increased maximal sample size n_{\max} if the observed effect is promising (decision $R = 1$) or to continue with a planned, original sample size n_0 if the observed effect is very promising (decision $R = 2$), $n_1 \leq n_0 < n_{\max}$. We summarize the decision R and sample size N based on the observed effect Y_1 as follows:

$$R = \begin{cases} 0 \\ 1 \\ 2 \end{cases} \quad \text{and } N = N(Y_1) = \begin{cases} n_f & \text{if } Y_1 \leq c_1 \\ n_{\max} & \text{if } c_1 < Y_1 \leq c_2 \\ n_0 & \text{if } Y_1 > c_2 \end{cases} \quad (1)$$

for some constants c_1, c_2 with $-\infty \leq c_1 < c_2 \leq \infty$. Usually $n_f = n_1$ is desired but sometimes this is not possible in practice and an ‘‘overrun’’ $n_f > n_1$ needs to be accepted. Note that a common two-stage group sequential design without sample size recalculation is an important special case of (1) with $n_f = n_0 = n_1$ (or allowing overrun with $n_f = n_0 \geq n_1$).

In the design defined through (1), we may calculate the bias of $\hat{\mu}_{WM,w}$ with $w(Y_1) = n_1/N = n_1/N(Y_1)$:

$$\begin{aligned} Cov[w(Y_1), Y_1] &= E[w(Y_1)Y_1] - E[w(Y_1)]E[Y_1] = n_1E[Y_1/N] - E[n_1/N]E[Y_1] \\ &= n_1\sigma_1 \left[-\frac{\phi(\frac{c_1-\mu}{\sigma_1})}{n_1} - \frac{\phi(\frac{c_2-\mu}{\sigma_1}) - \phi(\frac{c_1-\mu}{\sigma_1})}{n_{\max}} + \frac{\phi(\frac{c_2-\mu}{\sigma_1})}{n_0} \right]. \end{aligned}$$

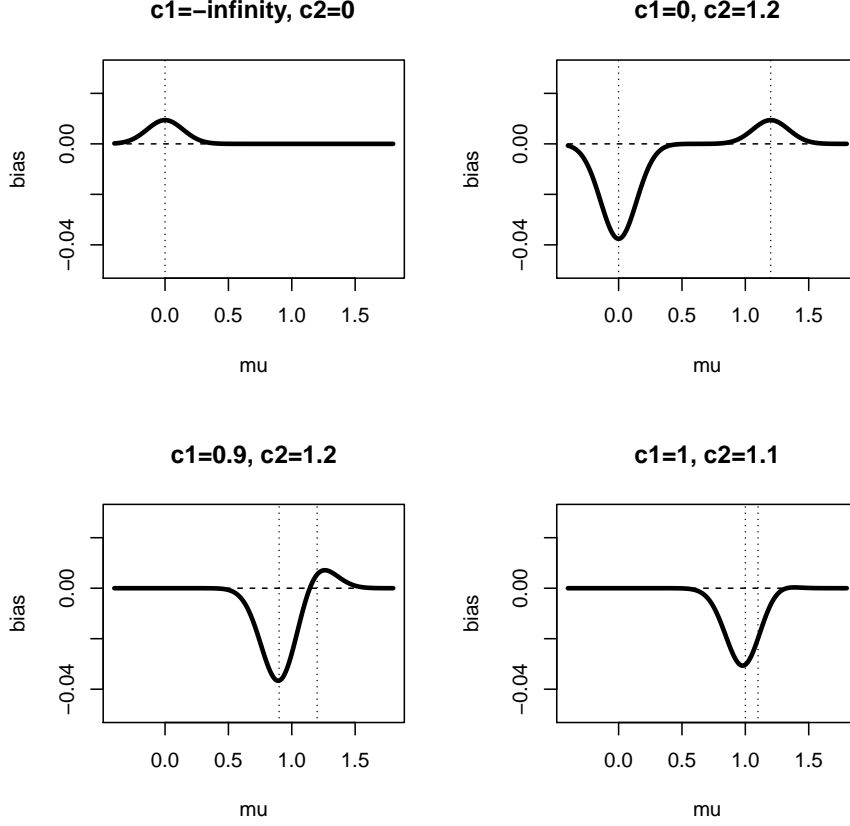


Figure 1: Bias of $\hat{\mu}_{WM,w}$ in dependence of μ for $n_1 = 50, n_0 = 100, n_{\max} = 150, \sigma = 1$ and for $(c_1, c_2) = (-\infty, 0)$ (upper left), $(c_1, c_2) = (0, 1.2)$ (upper right), $(c_1, c_2) = (0.9, 1.2)$ (lower left), $(c_1, c_2) = (1, 1.1)$ (lower right). The vertical dotted lines indicate c_1 and c_2 .

For example for $n_1 = 50, n_0 = 100, n_{\max} = 150$ and for $\sigma = 1$, the bias of $\hat{\mu}_{WM,w}$ in dependence of μ is shown in Figure 1 for some choices of (c_1, c_2) . The bias is negative around the value c_1 (where the sample size increases from $Y_1 < c_1$ to $Y_1 > c_1$) and is positive around c_2 (where the sample size decreases from $Y_1 < c_2$ to $Y_1 > c_2$). When c_1 and c_2 are very close (e.g. $c_1 = 1, c_2 = 1.1$, lower right panel), the positive bias can almost be hidden by the overlying negative bias.

2.2 Unconditionally unbiased estimates of the treatment effect

We can search for estimates $\hat{\mu}$ which are unbiased in the traditional, unconditional sense, i.e. $E[\hat{\mu}] = \mu$, by choosing appropriate weight functions $w(Y_1)$ for $\hat{\mu}_{WM,w}$. When the trial cannot stop at the interim, i.e. $c_1 = -\infty, n_{\max}, n_0 > n_1$, then any predetermined weights $w(Y_1) = w_1$ and $1 - w_1$,

$$\hat{\mu} = w_1 Y_1 + (1 - w_1) Y_2,$$

will ensure an unbiased estimate. According to [4], weights defined by $w_1 = 2n_1/(n_0 + n_{\max})$ may work fine. When early stopping is permitted, the reference [15] presents the unbiased estimate

$$\hat{\mu} = \frac{w_1 \sqrt{n_1} Y_1 + w_2 \sqrt{N_2} Y_2}{w_1 \sqrt{n_1} + w_2 \sqrt{N_2}},$$

for predetermined weights w_1, w_2 , suggested to be $w_1 = \sqrt{n_1/n_0}$ and $w_2 = \sqrt{1 - w_1^2}$. In the sequel, we will not restrict ourselves to predetermined weights and will permit that the weights change due to adaptations.

3 Conditional estimation after sample size recalculation

If the consequences of bias are different depending on the interim decision R , conditional estimation is sensible. In studies with an interim decision about futility stop only ($c_2 = \infty$, i.e. $R \in \{0, 1\}$), it is reasonable to require unbiasedness specifically if the trial is continued to Stage 2 ($R = 1$). Only if not stopped for futility, the results are used by regulatory agencies for decisions about licensing or if a Phase II trial is considered by the sponsor for decisions about continuation of the program to Phase III. Therefore, it is argued in [22] to require unbiasedness under the condition that the trial is continued, $E[\hat{\mu}|R = 1] = \mu$.

However, in our situation with decision rule (1), i.e. when based on an observed large effect the study is continued to a smaller sample size or stopped directly, the situation is a bit more complex. The interesting situations when good properties of the estimators are required are both $R = 1$ as well as $R = 2$. We investigate therefore the bias under the condition of R and show how to make the analysis unbiased in the conditional setting. As discussed before, the cases $R = 1$ and $R = 2$ are of main interest and we consider therefore these cases and not $R = 0$ in the sequel.

3.1 Uniformly Minimum Variance Unbiased Estimation

Consider first $R = 2$. Following the reference [13], which builds on [6], one may by means of the Rao-Blackwell theorem find an Unbiased Minimum Variance Estimate conditional on proceeding to stage two. There the situation is a drop-loser design, where the sample size of the second stage is given. In such a drop-loser design one would go forth to the second stage with selecting the best treatment and drop the others. In this article, however, we will derive results for only one treatment arm, thus without selection, and without control.

In contrast to the drop-the-looser situation with fixed sample sizes, we consider here a simple version of sample size re-calculation.

The conditional bias of the unconditional MLE $\hat{\mu}_{ML}$ equals

$$bias_{ML}(\mu, R) = \begin{cases} \sigma \frac{\sqrt{n_1}}{N} \frac{\phi(\bar{b}_1) - \phi(\bar{b}_2)}{\Phi(\bar{b}_1) - \Phi(\bar{b}_2)}, & R = 1, \\ \sigma \frac{\sqrt{n_1}}{N} \frac{\phi(\bar{b}_2)}{\Phi(\bar{b}_2)}, & R = 2, \end{cases} \quad (2)$$

where $\bar{b}_i = \bar{b}_i(\mu) = (\mu - c_i)/\sigma_1, i = 1, 2$, cf. (13.133) and (13.134) in [12]. However, since μ is unknown we cannot calculate this bias in practice. There is another way to construct an unbiased estimator which builds on the Rao-Blackwell theorem.

In the following Theorem, we present the Uniformly Minimum Variance Conditionally Unbiased Estimator (UMVCUE) and its conditional variance. The estimator consists of two terms: the unconditional MLE $\hat{\mu}_{ML}$ and a product of a standard deviation and a ratio involving ϕ and Φ , which is in the case of $R = 2$ equal to the inverse Mills ratio $\nu(x) = \phi(x)/\Phi(x)$. The unbiasedness is achieved by subtracting a term which closely resembles the bias presented above. The proof can be found in the Appendix

Theorem: The Uniformly Minimum Variance Conditionally Unbiased Estimator (UMVCUE) $\hat{\mu}_{RB}$ is given by

$$\hat{\mu}_{RB} = \begin{cases} g_1(\hat{\mu}_{ML}) = \hat{\mu}_{ML} - \sigma_B \frac{\Delta\phi}{\Delta\Phi}, & R = 1, \\ g_2(\hat{\mu}_{ML}) = \hat{\mu}_{ML} - \sigma_B \nu(z_2), & R = 2, \end{cases} \quad (3)$$

where $\Delta\phi = \phi(z_1) - \phi(z_2)$, $\Delta\Phi = \Phi(z_1) - \Phi(z_2)$, $z_i = z_i(\hat{\mu}_{ML}) = (\hat{\mu}_{ML} - c_i)/\sigma_A, i = 1, 2$, and $\sigma_A^2 = \sigma_1^4/(\sigma_1^2 + \sigma_2^2), \sigma_B^2 = \sigma_2^4/(\sigma_1^2 + \sigma_2^2)$.

The conditional variance is given by:

$$Var[\hat{\mu}_{RB}|R] \approx \begin{cases} g'_1(\mu_1)^2 Var[\hat{\mu}_{ML}|R = 1], & R = 1, \\ g'_2(\mu_2)^2 Var[\hat{\mu}_{ML}|R = 2], & R = 2, \end{cases} \quad (4)$$

where

$$g'_1(y) = 1 + \frac{\sigma_2^2}{\sigma_1^2} \frac{\{z_1(y)\phi(z_1(y)) + z_2(y)\phi(z_2(y))\} \Delta\Phi + \Delta\phi^2}{\Delta\Phi^2},$$

and

$$g'_2(y) = 1 + \frac{\sigma_2^2}{\sigma_1^2} \frac{z_2(y)\phi(z_2(y))\Phi(z_2(y)) + \phi(z_2(y))^2}{\Phi(z_2(y))^2},$$

and the variance factors in (4) can be found in the proof in formulae (13) and (14).

Note that the values of $\sigma_2, \sigma_A, \sigma_B$ are different for case $R = 1$ and $R = 2$ since they depend on N_2 .

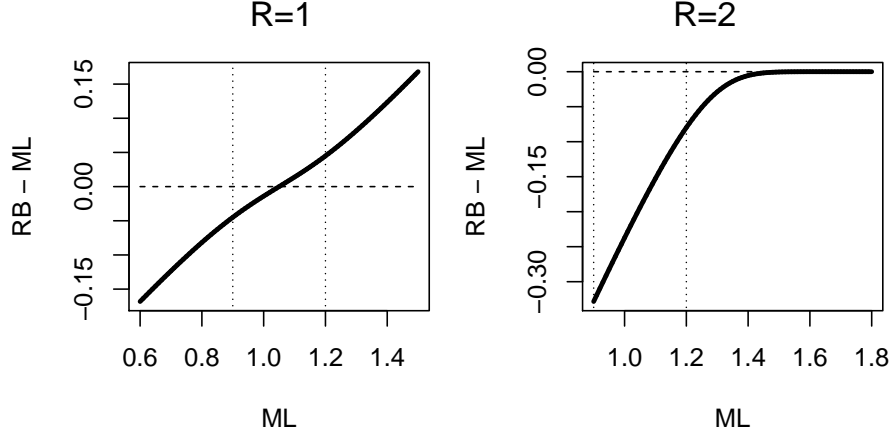


Figure 2: Difference between Rao-Blackwell (RB) and Maximum Likelihood (ML) estimator versus ML estimator for $R = 1$ (left) and $R = 2$ (right). Here $n_1 = 50, n_0 = 100, n_{max} = 150, c_1 = 0.9, c_2 = 1.2, \sigma = 1$. The vertical dotted lines indicate c_1 and c_2 .

To illustrate the difference between the Rao-Blackwell estimate in the above theorem and the Maximum Likelihood estimate, we show $\hat{\mu}_{RB} - \hat{\mu}_{ML}$ versus the MLE $\hat{\mu}_{ML}$ for $R = 1$ and $R = 2$, see Figure 2, with the specific values $n_1 = 50, n_0 = 100, n_{max} = 150, c_1 = 0.9, c_2 = 1.2$ and $\sigma = 1$. For example let us look at $R = 1$, i.e. when we have $0.9 < Y_1 < 1.2$ for the interim mean. Then the total sample size is raised from $n_0 = 100$ to $n_{max} = 150$. If the final mean, $\hat{\mu}_{ML}$, is still between 0.9 and 1.2, the MLE and the Rao-Blackwell differ at most 0.05. If $\hat{\mu}_{ML} > 1.2$, then the Rao-Blackwell estimate is larger than the MLE and if $\hat{\mu}_{ML} < 0.9$, then the Rao-Blackwell estimate is smaller than the MLE; this means in both cases that $\hat{\mu}_{RB}$ is shrunk towards Y_2 .

The fact that $\hat{\mu}_{RB} - \hat{\mu}_{ML}$ vs. $\hat{\mu}_{ML}$ is monotonically increasing for $R = 1$ and $R = 2$ implies that the Rao-Blackwell estimator overweights the second stage data compared to the first stage data. This is, if we define w_{RB} by $\hat{\mu}_{RB} = w_{RB}Y_1 + (1 - w_{RB})Y_2$, then $w_{RB} < n_1/(n_1 + N_2)$.

In Figure 2 for $R = 1$, $\hat{\mu}_{RB} - \hat{\mu}_{ML} > 0$ if and only if $\hat{\mu}_{ML} > 1.05$. Note that 1.05 is the mean of c_1 and c_2 . Further we recognize that $\hat{\mu}_{RB} - \hat{\mu}_{ML} < 0$ for $R = 2$. It follows from (3) that this observation is valid in general, see the following corollary:

Corollary: If $R = 1$, then

$$\hat{\mu}_{RB} \begin{cases} < \\ = \\ > \end{cases} \hat{\mu}_{ML} \text{ if } \hat{\mu}_{ML} \begin{cases} < \\ = \\ > \end{cases} \frac{c_1 + c_2}{2}.$$

If $R = 2$ then $\hat{\mu}_{RB} < \hat{\mu}_{ML}$.

Let us consider in general what happens for the case $R = 1$ if the interval $(c_1, c_2]$ becomes small. If $c_2 \rightarrow c_1$, $\hat{\mu}_{RB}$ will approach Y_2 . This follows as $\hat{\mu}_{RB} = E[Y_2|Y, Y_1 \in I] = E[Y_2|Y, Y_2 \in J]$, see the proof of the above Theorem.

3.2 Conditional Median Unbiased Estimation

The conditional median unbiased estimation comes naturally from interval estimation based on solving the equation

$$\int_{-\infty}^{\hat{\mu}_{ML}} f(y|\mu, R) dy = q, \quad (5)$$

with respect to μ for $q = 0.5$, cf. [14] and [25]. The solution μ of (5) is the conditional median unbiased estimator and will be denoted by $\hat{\mu}_{CMU}$. Note that interval estimation can be achieved by solving for q_1 and q_2 with $q_2 > 0.5 > q_1$ such that $q_2 - q_1 = l$ is the desired confidence level.

Take the case $R = 1$ first. As in the derivation of the Rao-Blackwell estimator we may show that the joint density for Y_1 and Y under the condition can be written as

$$f(y_1, y) = \tilde{k}(\mu) \phi\left(\frac{y - \mu}{\sigma_0}\right) \phi\left(\frac{y_1 - y}{\sigma_A}\right) \mathbf{1}\{c_1 < y_1 \leq c_2\}$$

with a function $\tilde{k}(\mu)$ independent of y_1 and y . Integrating out y_1 yields the conditional density:

$$f(y|\mu, R=1) = \frac{\Phi\left(\frac{c_2 - y}{\sigma_A}\right) - \Phi\left(\frac{c_1 - y}{\sigma_A}\right)}{\Phi\left(\frac{c_2 - \mu}{\sigma_1}\right) - \Phi\left(\frac{c_1 - \mu}{\sigma_1}\right)} \cdot \frac{\phi\left(\frac{y - \mu}{\sigma_0}\right)}{\sigma_0},$$

see also (13.133) in [12]. From this the case $R = 2$ follows:

$$f(y|\mu, R=2) = \frac{1 - \Phi\left(\frac{c_2 - y}{\sigma_A}\right)}{1 - \Phi\left(\frac{c_2 - \mu}{\sigma_1}\right)} \cdot \frac{\phi\left(\frac{y - \mu}{\sigma_0}\right)}{\sigma_0},$$

cf. equation (2) in [14].

We obtain a similar result when we compare $\hat{\mu}_{CMU}$ and $\hat{\mu}_{ML}$ as we have obtained for the comparison of $\hat{\mu}_{RB}$ and $\hat{\mu}_{ML}$ in Section 3.1.

Theorem: If $R = 1$, then

$$\hat{\mu}_{CMU} \begin{cases} < \\ = \\ > \end{cases} \hat{\mu}_{ML} \text{ if } \hat{\mu}_{ML} \begin{cases} < \\ = \\ > \end{cases} \frac{c_1 + c_2}{2}.$$

If $R = 2$ then $\hat{\mu}_{CMU} < \hat{\mu}_{ML}$.

Proof: Instead of (5), we can solve $\int_{-\infty}^{\hat{\mu}_{ML}} f(y|\mu, R) dy = \int_{\hat{\mu}_{ML}}^{\infty} f(y|\mu, R) dy$ (both integrals will be 0.5, then). Consequently, $\hat{\mu}_{CMU}$ is the solution μ of

$$\int_{-\infty}^{\hat{\mu}_{ML}} h(y) \cdot \phi\left(\frac{y - \mu}{\sigma_0}\right) dy = \int_{\hat{\mu}_{ML}}^{\infty} h(y) \cdot \phi\left(\frac{y - \mu}{\sigma_0}\right) dy \quad (6)$$

with $h(y) = \Phi\left(\frac{c_2 - y}{\sigma_A}\right) - \Phi\left(\frac{c_1 - y}{\sigma_A}\right)$ ($R = 1$) or $h(y) = 1 - \Phi\left(\frac{c_2 - y}{\sigma_A}\right)$ ($R = 2$).

$R = 1$: $h(y)$ is symmetric around $\frac{c_1 + c_2}{2}$. If $\hat{\mu}_{ML} = \frac{c_1 + c_2}{2}$, it follows from (6) that the normal density needs to have its mean at $\frac{c_1 + c_2}{2}$ as well. Therefore $\hat{\mu}_{CMU} = \hat{\mu}_{ML}$. If $\hat{\mu}_{ML} < \frac{c_1 + c_2}{2}$, the function $\phi\left(\frac{y - \mu}{\sigma_0}\right)$ needs to upweight the interval $(-\infty, \hat{\mu}_{ML}]$ compared to $[\hat{\mu}_{ML}, \infty)$ in order to fulfill (6). Hence $\hat{\mu}_{CMU} < \hat{\mu}_{ML}$ follows. If $\hat{\mu}_{ML} > \frac{c_1 + c_2}{2}$, we can show $\hat{\mu}_{CMU} > \hat{\mu}_{ML}$ similarly.

$R = 2$: $h(y)$ is monotonically increasing. Therefore the function $\phi\left(\frac{y - \mu}{\sigma_0}\right)$ needs to upweight the interval $(-\infty, \hat{\mu}_{ML}]$ in order to fulfill (6). We obtain $\hat{\mu}_{CMU} < \hat{\mu}_{ML}$. \square

Building on [14] one may derive an approximate bias formula for $\hat{\mu}_{CMU}$. Using equation (8) in [14] we may write

$$0.5 \approx F(\hat{\mu}_{ML}|\mu, R) + (\hat{\mu}_{CMU} - \mu) \frac{\partial F(\hat{\mu}_{ML}|\mu, R)}{\partial \mu}.$$

To denote the conditional expectation of $\hat{\mu}_{ML}$ use the notation $\mu_R = \mu + \text{bias}_{ML}(\mu, R)$, where the bias is specified in equation (2). This means that the bias approximately equals

$$\text{bias}_{CMU}(\mu, R) = E[\hat{\mu}_{CMU} - \mu|\mu, R] \approx \frac{0.5 - F(\mu_R|\mu, R)}{\frac{\partial F(\mu_R|\mu, R)}{\partial \mu}}. \quad (7)$$

For $R = 1$, the derivative equals

$$\frac{\partial F(x|\mu, R=1)}{\partial \mu} = \int_{-\infty}^x f(t|\mu, R=1) \left[\frac{t - \mu}{\sigma_0^2} + \frac{\phi(\frac{c_2 - \mu}{\sigma_1}) - \phi(\frac{c_1 - \mu}{\sigma_1})}{\sigma_1 \left\{ \Phi(\frac{c_2 - \mu}{\sigma_1}) - \Phi(\frac{c_1 - \mu}{\sigma_1}) \right\}} \right] dt. \quad (8)$$

Then evaluate at $x = \mu_1$, with $\mu_1 = \mu + \text{bias}_{ML}(\mu, 1)$, and similarly for $R = 2$. The Delta Method provides an estimate of the variance through

$$\text{Var}[\hat{\mu}_{CMU}|\mu, R] \approx \left\{ \frac{\partial}{\partial x} \frac{0.5 - F(x|\mu, R)}{\frac{\partial F(x|\mu, R)}{\partial \mu}} \Big|_{x=\mu_R} \right\}^2 \text{Var}[\hat{\mu}_{ML}|\mu, R].$$

In the scenarios tested however, approximation (7) appears rather inaccurate, see Section 4.

3.3 Conditional Maximum Likelihood Estimation

The reference [17] addresses conditional maximum likelihood estimation in a group sequential setting with efficacy stopping, and treats the problem with bivariate observations. Here, instead, we permit two kinds of continuation to the second stage, either with or without sample size adjustment, and, we derive an explicit bias correction formula.

We will derive the likelihood function given $R = 1$. For this we write $b_i = b_i(\mu) = \sqrt{n_1}(c_i - \mu)/\sigma$, $i = 1, 2$, and then,

$$P(R = 1) = P(c_1 < Y_1 \leq c_2) = \Phi(b_2(\mu)) - \Phi(b_1(\mu)). \quad (9)$$

The likelihood function is the product of a constrained multivariate distribution (for Stage 1) and an unconstrained (for Stage 2):

$$L_R(\mu) = \left\{ \prod_{j=1}^{n_1} \frac{1}{\sigma} \phi\left(\frac{x_j - \mu}{\sigma}\right) \right\} \cdot \frac{\mathbf{1}\{R=1\}}{P(R=1)} \times \prod_{j=n_1+1}^N \frac{1}{\sigma} \phi\left(\frac{x_j - \mu}{\sigma}\right).$$

Then the log likelihood becomes

$$\mathcal{L}_R(\mu) = -N \log(\sqrt{2\pi}\sigma) - \sum_{j=1}^N \frac{(x_j - \mu)^2}{2\sigma^2} - \log(\Phi(b_2(\mu)) - \Phi(b_1(\mu))).$$

Taking the derivative $\mathcal{L}'_R(\mu) = \frac{\partial \mathcal{L}_R(\mu)}{\partial \mu}$ with respect to μ and using $b'_i(\mu) = -\sqrt{n_1}/\sigma$ yields

$$\mathcal{L}'_R(\mu) = \sum_{j=1}^N \frac{x_j - \mu}{\sigma^2} + \frac{\sqrt{n_1}}{\sigma} \frac{\phi(b_2(\mu)) - \phi(b_1(\mu))}{\Phi(b_2(\mu)) - \Phi(b_1(\mu))}. \quad (10)$$

Note that the last term of $\mathcal{L}'_R(\mu)$ equals $-N \times \text{bias}_{ML}(\mu, R)/\sigma^2$, and, that the expected value of the first is $N \times \text{bias}_{ML}(\mu, R)/\sigma^2$. So, $E[\mathcal{L}'_R(\mu)|R] = 0$, cf. Lemma 6.1 of [16].

Solving the optimality condition

$$\sigma^2 \mathcal{L}'_R(\mu) = \sum_{j=1}^N (x_j - \mu) + \sqrt{n_1} \sigma \frac{\phi(b_2(\mu)) - \phi(b_1(\mu))}{\Phi(b_2(\mu)) - \Phi(b_1(\mu))} = 0$$

with respect to μ yields the conditional maximum likelihood estimate $\hat{\mu}_{CML}$.

An estimate of the variance of the estimator may be obtained through the observed information j

$$j(\hat{\mu}_{CML}) = \mathcal{L}''_R(\mu)|_{\mu=\hat{\mu}_{CML}} = -\frac{N}{\sigma^2} + \frac{b_2 v_2 - b_1 v_1}{s r} + \frac{(v_2 - v_1)^2}{r^2} \Big|_{\mu=\hat{\mu}_{CML}},$$

with $s = \sigma/\sqrt{n_1}$, $v_i = \phi(b_i(\mu))/s$, $b_i = b_i(\mu)$ and $r = P(R = 1) = \Phi(b_2) - \Phi(b_1)$. The variance estimate is $-j(\hat{\mu}_{CML})^{-1}$. In cases where asymptotics do not seem applicable a bootstrap approach

offers an alternative. Also, a likelihood based confidence interval employing a χ_1^2 approximation of $2(\mathcal{L}_R(\hat{\mu}_{CML}) - \mathcal{L}_R(\mu))$ is feasible.

As explained in [7] the bias of the MLE is related to the first three derivatives of the log likelihood. The derivative $\mathcal{L}_R'''(\mu)$ becomes

$$\frac{(b_2^2 - 1)v_2 - (b_1^2 - 1)v_1}{s^2 r} + \frac{(v_2 - v_1)(b_2 v_2 - b_1 v_1)}{s r^2} + \frac{2(v_2 - v_1)}{r^2} \left\{ \frac{b_2 v_2 - b_1 v_1}{s} + \frac{(v_2 - v_1)^2}{r} \right\}.$$

The formula (5.26) of [7] implies that the bias equals

$$bias_{CML}(\mu, R) = \{E(\mathcal{L}_R'''(\mu)|R)/2 + E(\mathcal{L}_R''(\mu)\mathcal{L}_R'(\mu)|R)\} / E(\mathcal{L}_R''(\mu)|R)^2 + \mathcal{O}(N^{-3/2}).$$

Please note that the second and third derivatives are independent of the sample. Further, as the expectation of the first derivative equals 0, the bias formula simplifies to

$$bias_{CML}(\mu, R) = \mathcal{L}_R'''(\mu) / \{2\mathcal{L}_R''(\mu)^2\} + \mathcal{O}(N^{-3/2}). \quad (11)$$

The above extends in an obvious way to any one-parameter MLE $\hat{\theta}(X)$, estimating an unknown parameter θ , noting that $P(R = 1) = P[c_1 < \hat{\theta}(X) \leq c_2]$ approximately equals $\Phi(-j(\hat{\theta})(c_2 - \hat{\theta})) - \Phi(-j(\hat{\theta})(c_1 - \hat{\theta}))$, cf. (9).

The case $R = 2$ follows when $c_2 \rightarrow \infty$ and then with c_2 replacing c_1 .

A bias corrected estimate $\hat{\mu}_{CMLc}$ comes from solving the equation

$$\hat{\mu}_{CML} = \mu + bias_{CML}(\mu, R) \quad (12)$$

with respect to μ , cf. equation (5.6.1) in [25]. We also tested the corrected estimate $\tilde{\mu}_{CMLc} = \hat{\mu}_{CML} - bias_{CML}(\hat{\mu}_{CML}, R)$, but the results were almost identical (data not shown).

Theorem: If $R = 1$, then

$$\hat{\mu}_{CML} \begin{cases} < \\ = \\ > \end{cases} \hat{\mu}_{ML} \text{ if } \hat{\mu}_{ML} \begin{cases} < \\ = \\ > \end{cases} \frac{c_1 + c_2}{2}.$$

If $R = 2$ then $\hat{\mu}_{CML} < \hat{\mu}_{ML}$.

Proof: Since the derivative in equation (10) equals zero at the optimum it follows that

$$\hat{\mu}_{ML} = \hat{\mu}_{CML} - \sigma \frac{\sqrt{n_1}}{N} \frac{\phi(b_2(\hat{\mu}_{CML})) - \phi(b_1(\hat{\mu}_{CML}))}{\Phi(b_2(\hat{\mu}_{CML})) - \Phi(b_1(\hat{\mu}_{CML}))},$$

from which the case $R = 1$ is immediate, noting that the denominator is positive, and, that: $sign(\phi(b_2(x)) - \phi(b_1(x))) = -sign(|b_2(x)| - |b_1(x)|)$. □

3.4 Difference between conditional estimators

We have seen empirically and proven algebraically that for all conditional estimators $\hat{\mu} \in \{\hat{\mu}_{RB}, \hat{\mu}_{CMU}, \hat{\mu}_{CML}\}$, the estimate is smaller than $\hat{\mu}_{ML}$ if $R = 1$ and $\hat{\mu}_{ML} < \frac{c_1 + c_2}{2}$ or if $R = 2$ and it is larger than $\hat{\mu}_{ML}$ if $R = 1$ and $\hat{\mu}_{ML} > \frac{c_1 + c_2}{2}$.

Based on numerical investigations, we investigate now the difference between the conditional estimators, $\hat{\mu}_{RB}, \hat{\mu}_{CMU}, \hat{\mu}_{CML}, \hat{\mu}_{CMLc}$. Figure 3 shows the difference for $n_1 = 50, n_0 = 100, n_{max} = 150, c_1 = 0.9, c_2 = 1.2, \sigma = 1$. The largest difference in this scenario between $\hat{\mu}_{CMU}$ and $\hat{\mu}_{RB}$ is for $R = 2$ and $\hat{\mu}_{ML} \approx 1.32$ and is -0.0066 . Compared to the difference between $\hat{\mu}_{RB}$ and $\hat{\mu}_{ML}$ seen before (Figure 2), this is a quite small difference. The difference between $\hat{\mu}_{CML}$ and $\hat{\mu}_{RB}$ is a little larger but still at most 0.01. We see that we can order the estimators in terms of how much they correct the ML: $\hat{\mu}_{RB}, \hat{\mu}_{CMU}, \hat{\mu}_{CML}$ (from smaller to larger absolute difference to ML). When we correct $\hat{\mu}_{CML}$ for the bias, the resulting $\hat{\mu}_{CMLc}$ is very similar to $\hat{\mu}_{RB}$.

When analysing numerically the difference between estimators for other scenarios, we have seen that this difference is often small. The estimators differ slightly more for smaller sample sizes and especially if $n_0 - n_1$ and $n_{max} - n_0$ become smaller.

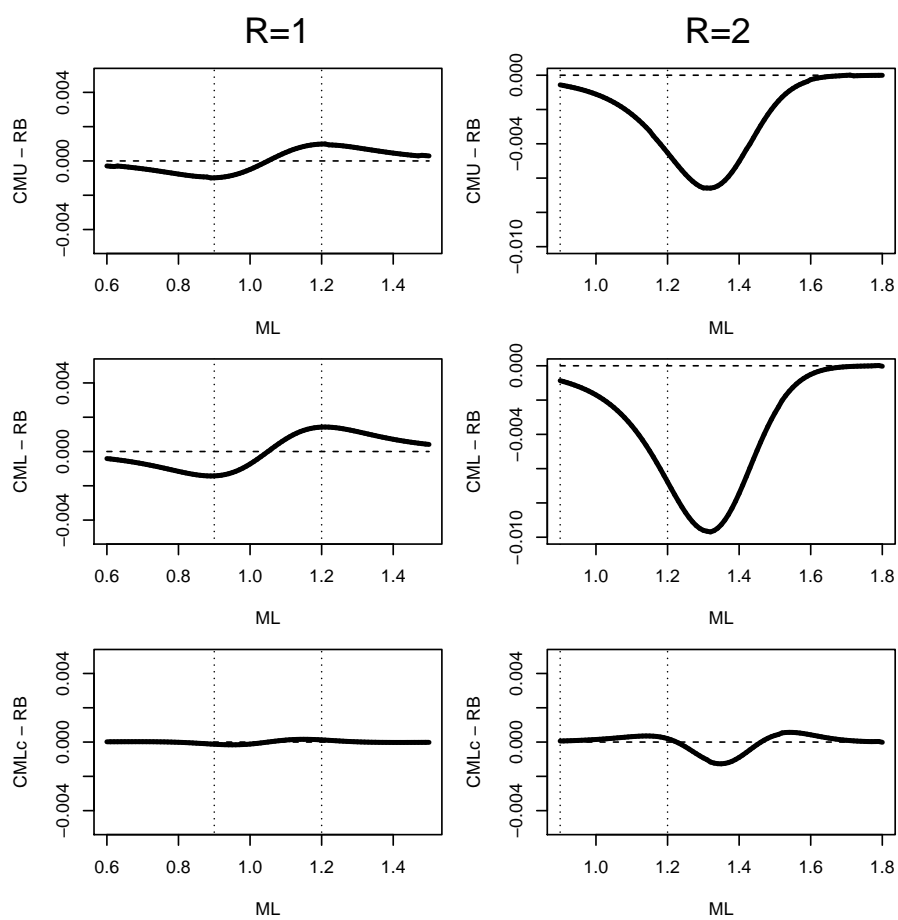


Figure 3: Difference between the estimators Conditional Median Unbiased (CMU), Conditional Maximum Likelihood (CML), corrected CML (CMLc) and Rao-Blackwell (RB), plotted versus Maximum Likelihood (ML) estimate. Here $n_1 = 50, n_0 = 100, n_{max} = 150, c_1 = 0.9, c_2 = 1.2, \sigma = 1$. The vertical dotted lines indicate c_1 and c_2 .

Table 1: Summary of bias and variation. Here $n_0 = n_1 + 50$, $n_{max} = 150$, $c_1 = 0.9$, $\sigma = 1$. 1000,000 simulations. The values for μ , n_1 , c_2 are chosen in 4 scenarios according to the values noted in the table.

μ	n_1	c_2	Case	Method	Bias	Var	MSE
1	50	1.2	R=1 N= 682108	RB	0.000	0.009	0.009
				CMU	-0.000	0.009	0.009
				CML	-0.000	0.009	0.009
				CMLc	0.000	0.009	0.009
				ML	0.011	0.005	0.005
			R=2 N= 78778	RB	0.001	0.017	0.017
				CMU	-0.002	0.017	0.017
				CML	-0.004	0.017	0.017
				CMLc	0.001	0.017	0.017
				ML	0.132	0.006	0.023
1.2	70	1.2	R=1 N= 494010	RB	-0.000	0.010	0.010
				CMU	0.002	0.010	0.010
				CML	0.003	0.010	0.010
				CMLc	0.000	0.010	0.010
				ML	-0.043	0.005	0.006
			R=2 N= 499857	RB	0.000	0.013	0.013
				CMU	-0.007	0.013	0.013
				CML	-0.010	0.013	0.013
				CMLc	-0.000	0.013	0.013
				ML	0.056	0.005	0.008
1.4	50	1.3	R=1 N= 239373	RB	0.000	0.009	0.009
				CMU	0.001	0.009	0.009
				CML	0.002	0.009	0.009
				CMLc	0.000	0.009	0.009
				ML	-0.061	0.005	0.009
			R=2 N= 760435	RB	0.000	0.013	0.013
				CMU	-0.005	0.013	0.013
				CML	-0.008	0.013	0.013
				CMLc	-0.001	0.013	0.013
				ML	0.029	0.008	0.009
0.9	50	1.2	R=1 N= 482187	RB	-0.000	0.009	0.009
				CMU	-0.001	0.009	0.009
				CML	-0.001	0.009	0.009
				CMLc	-0.000	0.009	0.009
				ML	0.035	0.005	0.006
			R=2 N= 16968	RB	-0.001	0.018	0.018
				CMU	-0.003	0.018	0.018
				CML	-0.005	0.017	0.017
				CMLc	-0.001	0.018	0.018
				ML	0.175	0.005	0.036

4 Simulation study

To assess the various methods we simulated normal data corresponding to a number scenarios and replicated a large number of times (1000,000). In each replication the outcome of Y_1 defined R and thereby decided the final design according to formula (1). Without loss of generality the value of σ was held fix at 1. We fixed also the value $c_1 = 0.9$ as the results depend only on the differences between the values for μ, c_1 and c_2 . We varied then μ over $\{0.9, 1, 1.2, 1.4\}$ and c_2 over $\{1.2, 1.3\}$. The maximal patient number (used when $R = 1$) was defined as $n_{max} = 150$ throughout, and n_0 was set to $n_1 + 50$, while n_1 varied over $\{50, 70\}$. These scenarios are presented in Table 1.

In Table 1 the methods are denoted as follows: Rao-Blackwell (RB), conditional median unbiased (CMU), conditional MLE (CML), bias corrected conditional MLE (CMLc), unconditional MLE (ML). In the simulations the CMLc used the correction $bias(\hat{\mu}|R)$ from equations (11) and (12). The tables provide the observed bias, variance and mean-square error (MSE) conditional on $R = 1$ and $R = 2$ as well as the number N of replications with $R = 1$ and $R = 2$, respectively.

Among the methods there is a high degree of concordance. From a practical point of view one could argue that MSE should take precedence over both bias and variance. As judged by MSE the methods RB, CMU, CML, and CMLc perform equally well. If we require strict conditional unbiasedness, then of course Rao-Blackwell is the only option.

However, one may argue that the naive unconditional MLE does quite well. In particular under $R = 1$, it performs best. If the true mean μ is large ($\geq c_2$, i.e. $R = 2$ is likely), the MLE performs also well under $R = 2$ (see second and third scenario in Table 1). In the cases when μ is small ($\leq c_1$, i.e. $R = 2$ is unlikely) but anyway $R = 2$ happened, the conditional MSE of the MLE is larger than for the other estimators (see first and last scenario in Table 1). Note that $R = 2$ happened in 8% and 2% of the simulations in these two scenarios, respectively.

The simulated bias in Table 1 can be compared with the theoretical biases stemming from (2), (7) and (11). The theoretical bias from (2) for ML coincides up to the third decimal with the simulated values in all cases. Also for CML, the simulated biases reflect the theoretically expected outcomes from (11) quite well with most absolute differences ≤ 0.001 and all ≤ 0.002 except the last scenario for $R = 2$ (simulated bias = -0.005 ; theoretical approximation = -0.001). In contrast, the approximation (7) appears more inaccurate. For example, the discrepancy in the second scenario in Table 1 ($\mu = 1.2, n_1 = 70, c_2 = 1.2$) for $R = 2$ is 0.011 (simulated bias = -0.007 ; theoretical approximation = 0.004). It is generally common that the simulated and theoretical bias have different signs.

5 Application in a schizophrenia trial

In a randomized, double-blind trial for clinically stable patients with schizophrenia, the treatments quetiapine and placebo were compared, [23]. The primary endpoint was time to first schizophrenic relapse analyzed by a Cox proportional hazards model. An interim analysis was conducted after 45 relapses observed (totally in both treatment groups) and the final analysis was scheduled after 90 relapses. (We ignore here for now the second interim which was preplanned – we will discuss this below.) We apply normal approximation to the survival analysis data and the total number of relapses are treated as sample size.

The trial was preplanned to be stopped when the interim two-sided p-value was < 0.004455 , i.e. stop for efficacy if $p < 0.004455$ to the advantage of quetiapine, and stop for futility if $p < 0.004455$ to the advantage of placebo. The trial had a fast recruitment speed and therefore, 61 relapses had been observed when the independent Data and Safety Monitoring Board's recommendation "stop" or "continue" from the interim analysis could be announced. The number of 61 relapses was not preplanned; nevertheless, we treat here the increase to 61 relapses as an unavoidable consequence of the recruitment process, like it would have been fixed in advance to have 61 relapses at this time.

When applying normal approximation (see Chapter 3.4 in [25] or Chapter 3.7 in [11]), we can formulate this situation with our sample size rule (1) with $n_1 = 45, n_0 = 61, n_{max} = 90$. The normal approximation implies that $-\log(\hat{H}R) \sim N(-\log(HR), 4/n)$ and therefore in accordance with our notation from Section 2, we have variance $\sigma^2 = 4$. We obtain $c_2 = \sqrt{\sigma^2/n_1}\Phi^{-1}(1 - 0.004455/2) = 0.848, c_1 = -c_2 = -0.848$.

According to [23], the observed hazard ratio was 0.16 after 45 relapses and 0.13 after 61 relapses. Taking $-\log(\hat{H}R)$, we have $Y_1 = 1.83$ in the interim and $\hat{\mu}_{ML} = 2.04$ in the final analysis. In order to

compute the Rao-Blackwell and conditional median unbiased estimator, we compute first $\sigma_1 = \sigma/\sqrt{n_1} = 2/\sqrt{45}$, $\sigma_2 = \sigma/\sqrt{N_2} = 1/2$, and $\sigma_A^2 = \sigma_1^4/(\sigma_1^2 + \sigma_2^2) = 0.0233$, $\sigma_B^2 = \sigma_2^4/(\sigma_1^2 + \sigma_2^2) = 0.1844$, leading to

$$\hat{\mu}_{RB} = \hat{\mu}_{ML} - \sigma_B \phi((\hat{\mu}_{ML} - c_2)/\sigma_A) / \Phi((\hat{\mu}_{ML} - c_2)/\sigma_A) = \hat{\mu}_{ML} - 10^{-14} \approx \hat{\mu}_{ML}.$$

Computing the other conditional estimates shows that in this case with a very clear result and $\hat{\mu}_{ML}$ far from c_2 , the estimates $\hat{\mu}_{CMU}$, $\hat{\mu}_{CML}$, $\hat{\mu}_{CMLc}$ are almost equal to $\hat{\mu}_{ML}$.

Let us assume that in contrast to the true trial results, the interim analysis would have just been over the stopping boundary with an estimated hazard ratio of 0.42 and that after 61 relapses the same estimate was obtained ($Y_1 = \hat{\mu}_{ML} = 0.87$). We obtain here $\hat{\mu}_{RB} = \hat{\mu}_{ML} - 0.304 = 0.566$ which corresponds to a Rao-Blackwell estimated hazard ratio of $\exp(-0.566) = 0.57$. The estimated hazard ratios with CMU, CML and CMLc are 0.59, 0.60, 0.57, respectively.

In this clinical trial, the result was a stop after interim, $R = 2$, and we condition on this outcome in the analysis. Therefore it does not matter for the conditional estimates how the design of the study would have been if $R = 1$ would have occurred. The above calculated estimates are therefore valid even for the real situation where a second interim analysis was preplanned.

6 Discussion

The naive MLE performs well in the simulations presented here. When comparing the unconditional MLE used in the conditional setting with the conditional estimators, the former had in many but not in all scenarios smaller MSE. However the unconditional MLE does not possess any optimality features in the conditional inference setting. Therefore it is worth while searching for alternatives that satisfy certain optimality criteria relevant for the conditional inference. The difference between the four conditional estimators was quite small in the scenarios considered. This was reflected also in terms of their bias (they had no bias or very small bias) and their variance which was similar. It is therefore not of main importance which of these to choose. The conditional Rao-Blackwell estimator has the advantage that it is unbiased by construction and has an explicit representation making computation simpler.

In this paper, we considered a sample size rule with general boundaries $c_1 < c_2$. In practice, one can choose these boundaries c_1 and c_2 to achieve a certain conditional or predictive power (see e.g. [11], ch. 10). Besides of considering these statistical criteria, we suggest to take non-statistical views into consideration: Usually an independent committee performs the interim analysis and recommends sample size for Stage 2 (or futility stop). The constants should then be chosen in order to mimic an anticipated way of reasoning of the committee (e.g. due to ethical arguments). Further, the decisions should be reasonable from an investment perspective. If these ethical and business arguments can be quantitatively specified, an optimisation of c_1 and c_2 might be an option. We refer to future research with regard to decision theoretic optimisation in this context and to a recent review article, [9], about decision theoretic methods in small studies.

It is straightforward to generalize our methods to a sample size rule with more steps, e.g. constant sample sizes between three cut-points c_1, c_2, c_3 . However, if an increasing number of cut-points are introduced, we will condition on more and more information from Stage 1. In the extreme case (for a continuous sample size recalculation formula) one would condition on the interim observation Y_1 . This would imply that the UMVCUE will be $\hat{\mu}_{RB} = Y_2$ as mentioned in Section 3.1. [3] discuss the fact that the information from Stage 1 is no longer used for the conditional estimate.

We point out that the methods described in this paper require the prespecification of the sample size recalculation rule. A further limitation of this elaboration is that we assumed known variance. In practice, variance is not known and one would use our methods with estimated variance. In this case the observed mean values are t - rather than normal-distributed. More generally, alternatives to our assumption of normally distributed data might be considered in future research. However, recall that we have pointed out ways to generalize to other one-parameter-families of distributions in Section 3.3. Further, if the sample sizes are large, the stagewise means are approximately normally distributed. Therefore we could apply our methods to the schizophrenia trial with a survival endpoint.

One may further extend our results concerning the one-arm model to two arms by use of profile likelihood. In a parallel group study the estimation problem may be expressed as follows, cf. [25]. The two independent series of observations obey laws $f(x, \mu_1, \eta)$ and $f(x, \mu_2, \eta)$, respectively, where η is a

vector of common nuisance parameters. The parameter of interest is the contrast $\theta = \frac{1}{2}(\mu_1 - \mu_2)$, while the level $\psi = \frac{1}{2}(\mu_1 + \mu_2)$ becomes one additional nuisance parameter.

Appendix

Proof of theorem in Section 3.1

Since Y_2 estimates μ without bias and Y is a complete and sufficient statistic for μ , we may invoke the Rao-Blackwell Theorem. That theorem requires calculation of $\hat{\mu}_{RB} = E[Y_2|Y, Y_1 \in I]$, where I denotes the outcome of the interim analysis.

We start with the case $R = 1$. Since $n_1 y_1 + (N - n_1) y_2 = Ny$, we have for given y that $y_1 \in (c_1, c_2]$ is equivalent to $y_2 \in J = \left[\frac{Ny - n_1 c_2}{N - n_1}, \frac{Ny - n_1 c_1}{N - n_1} \right)$. Hence $\hat{\mu}_{RB} = E[Y_2|Y, Y_1 \in I] = E[Y_2|Y, Y_2 \in J]$. To derive the joint distribution of Y and Y_2 , we start with the joint density of Y_1 and Y_2 under the conditioning which is proportional to

$$\frac{1}{\sigma_1} \phi\left(\frac{y_1 - \mu}{\sigma_1}\right) \frac{1}{\sigma_2} \phi\left(\frac{y_2 - \mu}{\sigma_2}\right) \mathbf{1}\{y_2 \in J\} \mathbf{1}\{n_1 y_1 + (N - n_1) y_2 = Ny\},$$

where $\mathbf{1}\{\cdot\}$ is the indicator function. When we multiply the part before the indicator functions with $2\pi\sigma_1\sigma_2$, take the logarithm and multiply with $-1/2$, we get:

$$\left(\frac{y}{\sigma_0^2} - \frac{y_2}{\sigma_2^2} - \frac{\mu}{\sigma_1^2}\right)^2 \sigma_1^2 + \left(\frac{y_2 - \mu}{\sigma_2}\right)^2$$

since $y/\sigma_0^2 = y_1/\sigma_1^2 + y_2/\sigma_2^2$. It can be shown that this is equal to

$$\left(\frac{y - \mu}{\sigma_0}\right)^2 + \left(\frac{y_2 - y}{\sigma_2 \sigma_0}\right)^2 \sigma_1^2 + k(\mu)$$

for a function $k(\mu)$ independent of y_2 and y . Hence, we obtain that the joint density for Y_2 and Y under the condition can be written as

$$f(y_2, y) = \tilde{k}(\mu) \phi\left(\frac{y - \mu}{\sigma_0}\right) \phi\left(\frac{y - y_2}{\sigma_B}\right) \mathbf{1}\{y_2 \in J\}$$

with $\tilde{k}(\mu)$ independent of y_2 and y and with $\sigma_B = \sigma_0 \sigma_2 / \sigma_1 = \sigma_2^2 / \sqrt{\sigma_1^2 + \sigma_2^2}$.

The above joint density implies that $E[Y_2|Y, Y_2 \in J]$ equals to a truncated normal density, where the original normal distribution before truncation has mean y and variance σ_B^2 . According to (13.134) in [12], we get

$$\hat{\mu}_{RB} = y - \sigma_B \frac{\phi\left(\left(\frac{Ny - n_1 c_1}{N - n_1} - y\right) / \sigma_B\right) - \phi\left(\left(\frac{Ny - n_1 c_2}{N - n_1} - y\right) / \sigma_B\right)}{\Phi\left(\left(\frac{Ny - n_1 c_1}{N - n_1} - y\right) / \sigma_B\right) - \Phi\left(\left(\frac{Ny - n_1 c_2}{N - n_1} - y\right) / \sigma_B\right)}.$$

Let $z_i = \left(\frac{Ny - n_1 c_i}{N - n_1} - y\right) / \sigma_B = \frac{n_1}{N - n_1} (y - c_i) / \sigma_B = (y - c_i) / \sigma_A$, $i = 1, 2$, with $\sigma_A = \sigma_B \frac{N - n_1}{n_1} = \sigma_1 \sigma_0 / \sigma_2 = \sigma_1^2 / \sqrt{\sigma_1^2 + \sigma_2^2}$. Then we can write

$$\hat{\mu}_{RB} = y - \sigma_B \frac{\phi(z_1) - \phi(z_2)}{\Phi(z_1) - \Phi(z_2)}.$$

The case $R = 2$ can be treated as special case of $R = 1$ by setting $c_2 = \infty$ and replacing c_1 by c_2 leading to $\hat{\mu}_{RB} = y - \sigma_B \frac{\phi(z_2)}{\Phi(z_2)}$. See also [13] for the case $R = 2$.

The conditional variance of $\hat{\mu}_{RB}$ may be obtained through the Delta Method, cf. [14]. For $R = 1$, the approximate conditional variance follows from application of that method by writing $\hat{\mu}_{RB}$ as a function of y ,

$$g_1(y) = y - \sigma_B \frac{\phi(z_1(y)) - \phi(z_2(y))}{\Phi(z_1(y)) - \Phi(z_2(y))}.$$

Then

$$g_1'(y) = 1 + \sigma_B/\sigma_A \frac{\{z_1\phi(z_1) + z_2\phi(z_2)\}(\Phi(z_1) - \Phi(z_2)) + (\phi(z_1) - \phi(z_2))^2}{(\Phi(z_1) - \Phi(z_2))^2}.$$

From (13.135) in [12] one may prove

$$\text{Var}[\hat{\mu}_{ML}|R=1] = \frac{1}{N^2} (\text{Var}[n_1Y_1 + n_2Y_2|R=1]) = \frac{1}{N^2} (n_1^2 \text{Var}[Y_1|R=1] + n_2\sigma^2), \quad (13)$$

where

$$\text{Var}[Y_1|R=1] = \frac{\sigma^2}{n_1} \left[1 + \frac{b_1\phi(b_1) - b_2\phi(b_2)}{\Phi(b_2) - \Phi(b_1)} - \left\{ \frac{\phi(b_1) - \phi(b_2)}{\Phi(b_2) - \Phi(b_1)} \right\}^2 \right]$$

for $b_i = \sqrt{n_1}(c_i - \mu)$, $i = 1, 2$.

For $R = 2$, we may write the Rao-Blackwell estimator as $g_2(\hat{\mu}_{ML})$ with $g_2(y) = y - \sigma_B\nu(z_2(y))$, $z_2(y) = (y - c_2)/\sigma_A$. Invoking the Delta Method we have

$$\text{Var}[g_2(\hat{\mu}_{ML})|R=2] \approx g_2'(\mu_2)^2 \text{Var}[\hat{\mu}_{ML}|R=2]$$

and $g_2'(y) = 1 - \nu'(z_2(y))z_2'(y)$, $\nu'(t) = -\phi(t)(t\Phi(t) + \phi(t))/\Phi^2(t)$ with $t = z_2(y)$, and, $z_2'(y) = 1/\sigma_A$. As in the case $R = 1$ we may derive

$$\text{Var}[\hat{\mu}_{ML}|R=2] = \frac{1}{N^2} (\text{Var}[n_1Y_1 + n_2Y_2|R=2]) = \frac{1}{N^2} (n_1^2 \text{Var}[Y_1|R=2] + n_2\sigma^2), \quad (14)$$

where

$$\text{Var}[Y_1|R=2] = \frac{\sigma^2}{n_1} \left[1 + \frac{b_2\phi(b_2)}{1 - \Phi(b_2)} - \left\{ \frac{\phi(b_2)}{1 - \Phi(b_2)} \right\}^2 \right] \cdot \square$$

References

- [1] BRANNATH, W., KÖNIG, F., AND BAUER, P. Estimation in flexible two stage designs. *Statistics in Medicine* 25, 19 (2006), 3366–3381.
- [2] BROBERG, P. Sample size re-assessment leading to a raised sample size does not inflate type I error rate under mild conditions. *BMC Medical Research Methodology* 13, 1 (2013), 94.
- [3] BURMAN, C., AND SONESSON, C. Are flexible designs sound? *Biometrics* 62 (2006), 664–683.
- [4] CHANG, M. *Adaptive Design Theory and Implementation Using SAS and R*. Chapman and Hall/CRC, Boca Raton, 2007.
- [5] CHEN, Y. H. J., DEMETS, D. L., AND GORDON LAN, K. K. Increasing the sample size when the unblinded interim result is promising. *Statistics in Medicine* 23, 7 (2004), 1023–1038.
- [6] COHEN, A., AND SACKROWITZ, H. B. Two stage conditionally unbiased estimators of the selected mean. *Statistics & Probability Letters* 8, 3 (1989), 273–278.
- [7] COX, D., AND BARNDORFF-NIELSEN, O. *Inference and Asymptotics*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Taylor & Francis, London, 1994.
- [8] CUFFE, R. L., LAWRENCE, D., STONE, A., AND VANDEMEULEBROECKE, M. When is a seamless study desirable? case studies from different pharmaceutical sponsors. *Pharmaceutical Statistics* 13, 4 (2014), 229–237.
- [9] HEE, S., HAMBORG, T., DAY, S., MADAN, J., MILLER, F., POSCH, M., ZOHAR, S., AND STALLARD, N. Decision theoretic designs for small trials and pilot studies: a review. *Statistical Methods in Medical Research to appear* (2015).
- [10] IASONOS, A., AND O’QUIGLEY, J. Adaptive dose-finding studies: A review of model-guided phase I clinical trials. *Journal of Clinical Oncology* 32, 23 (2014), 2505–2511.
- [11] JENNISON, C., AND TURNBULL, B. *Group sequential methods with applications to clinical trials*. Chapman & Hall/CRC, Boca Raton, 2000.
- [12] JOHNSON, N., KOTZ, S., AND BALAKRISHNAN, N. *Continuous Univariate Distributions, Volume 1*. Wiley, New York, 1994, ch. 13.10.1.
- [13] KIMANI, P. K., TODD, S., AND STALLARD, N. Conditionally unbiased estimation in phase II/III clinical trials with early stopping for futility. *Statistics in Medicine* 32, 17 (2013), 2893–2910.

- [14] KOOPMEINERS, J. S., FENG, Z., AND PEPE, M. S. Conditional estimation after a two-stage diagnostic biomarker study that allows early termination for futility. *Statistics in Medicine* 31, 5 (2012), 420–435.
- [15] LAWRENCE, J., AND HUNG, H. J. Estimation and confidence intervals after adjusting the maximum information. *Biometrical Journal* 45, 2 (2003), 143–152.
- [16] LEHMANN, E. *Theory of Point Estimation*, first ed. Wiley, New York, 1983.
- [17] LIU, A., TROENDLE, J. F., YU, K. F., AND YUAN, V. W. Conditional maximum likelihood estimation following a group sequential test. *Biometrical Journal* 46, 6 (2004), 760–768.
- [18] MEHTA, C. R., AND POCOCK, S. J. Adaptive increase in sample size when interim results are promising: A practical guide with examples. *Statistics in Medicine* 30, 28 (2011), 3267–3284.
- [19] MILLER, F., BJÖRNSSON, M., SVENSSON, O., AND KARLSTEN, R. Experiences with an adaptive design for a dose-finding study in patients with osteoarthritis. *Contemporary Clinical Trials* 37, 2 (2014), 189–199.
- [20] MÜLLER, H.-H., AND SCHÄFER, H. Adaptive group sequential designs for clinical trials: Combining the advantages of adaptive and of classical group sequential approaches. *Biometrics* 57 (2001), 886–891.
- [21] MÜLLER, H.-H., AND SCHÄFER, H. A general statistical principle for changing a design any time during the course of a trial. *Statistics in Medicine* 23 (2004), 2497–2508.
- [22] PEPE, M. S., FENG, Z., LONGTON, G., AND KOOPMEINERS, J. Conditional estimation of sensitivity and specificity from a phase 2 biomarker study allowing early termination for futility. *Statistics in Medicine* 28 (2009), 762–779.
- [23] PEUSKENS, J., TRIVEDI, J., MALYAROV, S., BRECHER, M., SVENSSON, O., MILLER, F., PERS-SON, I., AND MEULIEN, D. ON BEHALF OF THE STUDY D1444C00004 INVESTIGATORS. Prevention of schizophrenia relapse with extended release quetiapine fumarate dosed once daily: A randomized, placebo-controlled trial in clinically stable patients. *Psychiatry* 4 (2007), 34–50.
- [24] POSCH, M., BAUER, P., AND BRANNATH, W. Issues in designing flexible trials. *Statistics in Medicine* 22 (2003), 953–969.
- [25] WHITEHEAD, J. *The design and analysis of sequential clinical trials*. Ellis Horwood series in mathematics and its applications. Ellis Horwood Limited, Chichester, 1992.